Simulation study of the collapse of linear and ring homopolymers

Jianpeng Ma and John E. Straub
Department of Chemistry, Boston University, Boston, Massachusetts 02215

Eugene I. Shakhnovich
Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

The thermodynamic and kinetic properties of Lennard-Jones homopolymers are studied by computer simulation of the homopolymer at constant temperature and during collapse following an abrupt temperature drop. The homopolymer dynamics is simulated using both the Gaussian phase packet dynamics algorithm, which provides an approximate solution of the Liouville equation for the time dependent classical density distribution, and conventional molecular dynamics. Numerical simulations of both linear and ring polymers are analyzed and compared with theoretical predictions. The results support the idea that homopolymer collapse is a two-stage process. © *1995 American Institute of Physics.*

I. INTRODUCTION

Understanding the mechanism of protein folding remains a central problem of molecular biology. The problem of how to determine the final unique globular conformation and the dynamic folding pathway continues to be the focus of many researchers.[1] Considerable progress has been made in both experimental[2] and theoretical[3] studies. Recent experimental studies[4] and computer simulation[5-11] studies strongly suggest that protein folding is a multistage process, the first stage being the collapse of the molecule to a compact conformation, followed by rearrangement to acquire the native fold. The first stage of folding, compaction, was studied in detail recently by Socci and Onuchic[9] using Monte Carlo simulations of a short, 27 monomer lattice model protein. These authors showed that the rate of this burst compactization stage is sequence-independent, which makes it similar to the collapse of a homopolymer. Therefore it is reasonable to believe that insights into the character and mechanism of the fastest stage of protein folding may be gained from a detailed analysis of the thermodynamics and dynamics of collapse transitions in a simple homopolymer. Another reason to study the collapse transition in homopolymers is that it may help to test various analytical approaches and assess the validity of different approximations made in theoretical studies.[12,13] In the present paper, we restrict our attention to a Lennard-Jones homopolymer of N monomers.

As one of the most important phenomena in polymer physics, the coil-globule phase transition vividly illustrates many fundamental properties of polymer systems. The pioneering research on this transition was stimulated by the puzzle of protein denaturation, in which the protein molecule undergoes a cooperative, sharp transition accompanied by changes of many of its physicochemical characteristics, including loss of biological activity. The approach to the coil-globular transition based on a generalization of Flory theory was proposed by Ptitsyn and Eisner in 1964.[12] The fundamental theoretical approach to the study of the coil-globule transition in polymer systems was proposed by Lifshitz[14] who pointed out that the coil and the globule are two different macroscopic phase states of the polymer chain. These two phases differ primarily in their fluctuation regimes. The radius of correlation of density in the globule is small compared to its size; in the coil state it is of the order of its size. The phase transition between the coil and globule state was predicted[14] to be a weak first order transition with a latent heat and density jump per particle which vanish at the transition point as the size of the molecule increases. The simplest macromolecular model which manifests the fundamental properties addressed in Lifshitz' theory is the so-called standard Gaussian model. In this model, the polymer molecule is represented as N beads connected by a flexible string; thus the Gaussian model is also called a ''beads'' model.

A general scenario for the kinetics of the collapse of a flexible polymer coil has been presented by de Gennes.[15] An ''expanding sausage model'' was used to describe the collapse of the polymer molecule. His scenario predicts that the polymer chain first collapses to a sausage-shaped object where the subunits of the chain, called ''blobs,'' tend to stick together. This corresponds to the relaxation of short length scale properties. In the second stage of collapse, the ''sausage'' grows with the hydrodynamic friction until the molecule reaches the final spherical form. During the collapse of the polymer, the density is predicted to increase (and reach its equilibrium value) much faster than the shape of the molecule relaxes to equilibrium.

A more detailed scenario of the kinetics of homopolymer collapse was presented by Grosberg, Nechaev, and Shakhnovich (GNS).[16] They proposed that the kinetics of collapse of a polymer coil following an abrupt decrease in temperature is a two-stage process. The first stage leads to a so-called ''crumpled globule'' where the monomers proximal along the chain are most likely to be proximal in space. The crumpled globule is predicted to be formed through a self-similar process so that its features persist on all length scales. This makes a crumpled globule a fractal object. On a length scale larger than N_e the chain domains of the fractal globule are segregated from each other in space. (N_e is the well-known parameter of the reptation model.[13,17]) For the crumpled

globule, the density inside the "blobs" is close to the final globular density while the "blobs" are still spatially segregated.

The second stage of homopolymer collapse in the GNS scenario is chain quasi-knotting in which the crumpled globule continues to shrink through topological relaxation to the final equilibrium state. The key difference between the crumpled and equilibrium globules is that in the crumpled globule monomers which are neighbors along the chain are more likely to be neighbors in space, while in the equilibrium globule monomers distant along the chains are more likely to be in contact. Thus, in a linear chain the ends play a crucial role in this process, which is proposed to follow a reptation-type mechanism. Not surprisingly, the characteristic time of this process scales with the number of monomers N as N^3. Naturally, a ring polymer would not have this second relaxation stage; for such polymers the crumpled globule state is expected to be the natural consequence of compactization. Compact ring polymers are predicted to be in the crumpled globule state for an indefinitely long time.

A test of these theoretical predictions in a real or numerical experiment is important to estimate the adequacy of the concept of a crumpled globule. In this paper, we present a computer simulation study of homopolymer collapse and compare our results with the predictions of the GNS scenario. In our model homopolymer, the interaction potential between any two beads is a typical Lennard-Jones 12-6 potential form. We study in parallel the behavior of ring and linear polymers because, as was stated above, they are predicted to differ in the character of their equilibrium compact state.

We employed the conventional molecular dynamics (MD) method and an approximate numerical simulation technique, Gaussian phase packet (GPP) dynamics, developed recently by Ma, Hsu, and Straub[18] as an alternative to the normal MD method, as our simulation tools. The GPP method is based on an approximate solution of the classical Liouville equation using a product of single particle Gaussian phase packets to represent the phase space density distribution. The parameters of the density distribution obey variationally optimized equations of motion. As such, one can replace the calculation of a single trajectory in normal MD, which explores a line of points in phase space, with a continuous distribution of trajectories, which visit a much larger volume of phase space. The GPP method can be generally regarded as an approximate enhanced sampling algorithm. For applications to Lennard-Jones clusters and fluids, the phase space sampling statistics are improved over those of conventional molecular dynamics simulation. However, the results are approximate while those of MD are in principle exact. The parallel comparison study of the MD and GPP methods demonstrates the advantages and shortcomings of the GPP method.

In this paper, after a brief description of the statistical mechanics background in section II, we outline the formalism of GPP dynamics at constant temperature and the corresponding temperature jump strategy. In section III, we discuss the simulation results for both the thermodynamic and kinetic properties of the linear and ring Lennard-Jones homopolymers using the MD and GPP methods. We focus on the comparison of simulation data and theoretical predictions for the collapse kinetics. A brief summary and discussion is provided in section IV.

II. METHOD

In this section, we begin with a discussion of the Gaussian phase packet (GPP) approximation to the classical density distribution. We then define the equations of motion of the classical density distribution in the GPP approximation at constant temperature.

A. Statistical mechanics background of the GPP method

The standard canonical ensemble average of a physical quantity $A(\mathbf{r},\mathbf{p})$ takes the form[19]

$$\langle A \rangle = \int d^d \mathbf{r} \int d^d \mathbf{p} \rho(\mathbf{r},\mathbf{p},t) A(\mathbf{r},\mathbf{p}), \tag{1}$$

where $\rho(\mathbf{r},\mathbf{p},t)$ is the phase space density distribution, the time evolution of which is described by the Liouville equation[20,21]

$$\frac{\partial}{\partial t} \rho(\mathbf{r},\mathbf{p},t) = -\mathscr{L}_0 \rho(\mathbf{r},\mathbf{p},t), \tag{2}$$

where $\mathscr{L}_0$ is the Liouville operator

$$\mathscr{L}_0 = \frac{\mathbf{p}}{M} \cdot \nabla_{\mathbf{r}} + \mathbf{F}(\mathbf{r}) \cdot \nabla_{\mathbf{p}} \tag{3}$$

and where $\mathbf{F}(\mathbf{r})$ is the force and M the mass. $\mathbf{F}(\mathbf{r})$, $\mathbf{r}$, and $\mathbf{p}$ are d-dimensional vectors. In the Gaussian phase packet approximation, the phase space density distribution function for each particle in d dimensions is a single spherically symmetric Gaussian function

$$\rho(\mathbf{r},\mathbf{p},t) = \left[\frac{\sqrt{1-\alpha^2}}{2\pi\sigma_a\sigma_b} \right]^d \exp\left[-\frac{1}{2}\left(\frac{\mathbf{r}-\mathbf{r}_0}{\sigma_a}\right)^2 - \frac{1}{2}\left(\frac{\mathbf{p}-\mathbf{p}_0}{\sigma_b}\right)^2 \right.$$
$$\left. - \alpha\left(\frac{\mathbf{r}-\mathbf{r}_0}{\sigma_a}\right)\left(\frac{\mathbf{p}-\mathbf{p}_0}{\sigma_b}\right) \right]. \tag{4}$$

$\mathbf{r}_0$, $\mathbf{p}_0$, σ_a, σ_b, and α, which depend explicitly on time, must be calculated in order to fully define $\rho(\mathbf{r},\mathbf{p},t)$. The equations of motion can be derived for both constant energy and constant temperature dynamics.[18]

B. Gaussian phase packets at constant temperature

The constant temperature constraint can be applied using Gauss' principle.[22] We write the generalized Liouvillian

$$\mathscr{L} = \frac{\mathbf{p}}{M} \cdot \nabla_{\mathbf{r}} + (\mathbf{F}(\mathbf{r}) - \gamma\mathbf{p}) \cdot \nabla_{\mathbf{p}} = \mathscr{L}_0 - \gamma\mathbf{p} \cdot \nabla_{\mathbf{p}}, \tag{5}$$

where γ is determined by the general temperature (T) constraint equation

$$\frac{dT}{dt} = \frac{1}{dk_B M} \frac{d\langle \mathbf{p}^2 \rangle}{dt} = -\eta(t). \tag{6}$$

If η is zero, the temperature is held rigorously constant. For η equal to a constant greater than zero, the temperature will decrease linearly in time.

The resulting equations of motion for the Gaussian phase packet are[18]

$$\dot{\mathbf{r}}_0 = \frac{\mathbf{p}_0}{M}, \quad \dot{\mathbf{p}}_0 = -\nabla_{\mathbf{r}_0}\langle V\rangle - \gamma\mathbf{p}_0,$$

$$\dot{M}_{2,0} = \frac{2}{M}M_{1,1}, \quad \dot{M}_{1,1} = \frac{1}{M}M_{0,2} - \frac{1}{d}M_{2,0}\nabla_{\mathbf{r}_0}^2\langle V\rangle - \gamma M_{1,1},$$

$$\dot{M}_{0,2} = -\frac{2}{d}M_{1,1}\nabla_{\mathbf{r}_0}^2\langle V\rangle - 2\gamma M_{0,2}, \tag{7}$$

where the $M_{n,k}$ are the moments of the distribution defined as $M_{n,k} = \langle(\mathbf{r}-\mathbf{r}_0)^n(\mathbf{p}-\mathbf{p}_0)^k\rangle$. Note that $\langle V\rangle$ is the bare potential energy averaged over the phase space distribution

$$\langle V\rangle \equiv \int d^d\mathbf{r}\int d^d\mathbf{p}\,\rho(\mathbf{r},\mathbf{p},t)V(\mathbf{r}). \tag{8}$$

The value of γ is determined by the constraint equation

$$\gamma = \frac{dk_BM\eta/2 + \mathbf{p}_0\cdot\mathbf{F}_0 - M_{1,1}\nabla_{r_0}^2\langle V\rangle/d}{\mathbf{p}_0 + M_{0,2}}, \tag{9}$$

where the value of $\eta(t)$ determines the cooling schedule. We have used the fact that $\langle\mathbf{p}^2\rangle = \mathbf{p}_0^2 + M_{0,2}$. While the discussion above is focussed on a single particle system in d dimensions, generalization of these equations to an N-body system is straightforward when the N-body density distribution is approximated:

$$\rho(\mathbf{r}^N,\mathbf{p}^N,t) = \prod_{k=1}^{N}\rho(\mathbf{r}_k,\mathbf{p}_k,t) \tag{10}$$

as a product of single particle density distributions.[18,23]

III. APPLICATION TO LJ HOMOPOLYMERS

We begin with a general description of the numerical simulation strategy. This is followed by the results for the thermodynamic and dynamic properties of both linear and ring polymers.

A. Numerical simulations

In our simulation, the polymer chains were modeled using (1) a relatively soft harmonic bond between neighboring sites

$$V'_{ij} = \tfrac{1}{2}\kappa(|\mathbf{r}_{ij}| - l_b)^2, \tag{11}$$

where $l_b = 2^{1/6}$ is the equilibrium bond length and $\kappa = 10$ is the bond force constant and (2) a Lennard-Jones pair potential acting between all pairs of sites. The total potential energy is a sum over pair potentials of the form

$$V(r_{ij}) = 4(r_{ij}^{-12} - r_{ij}^{-6}), \tag{12}$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. The harmonic forces between adjacent sites determine the connectivity while the bond lengths are allowed to fluctuate. A detailed discussion of the bond potential is given in the Appendix.

The effective potential [Eq. (8)] corresponding to the Lennard-Jones potential cannot be calculated in a closed form for the GPP representation of the density distribution. We chose to fit the Lennard-Jones 6-12 potential to a sum of four Gaussians.[24] The parameters (a_k, b_k) for the potential fit of $V(\mathbf{r}) = \Sigma_k a_k\exp(-b_k\mathbf{r}^2/2)$ given in Lennard-Jones reduced units are $(8.467\,067\times10^5, 30.928\,81)$, $(2.713\,651\times10^3, 14.963\,75)$, $(-0.7154420, 1.279242)$, and $(-9.699\,172, 3.700\,745)$. These parameters were derived from a fit to the force and were found to provide a good representation of the Lennard-Jones 12-6 potential and the corresponding force. Lennard-Jones reduced units are used throughout the study.

All of the MD and GPP simulations employed the following protocol for both linear and ring polymers.

(1) The initial polymer configurations were chosen randomly in three-dimensional space with the bond lengths set equal to the equilibrium length $a = 2^{1/6}$ in Lennard-Jones reduced units. The initial random conformations of ring polymers were generated by HyperChem on a Silicon Graphic workstation. Typically, the initial bond lengths of the ring polymers were not exactly equal to the equilibrium distance. Any small deviations relaxed during the equilibration.

(2) The equations of motion [Eqs. (7) and (9)] were integrated using the Bulirsch-Stoer method[25] with an internal relative error tolerance of 10^{-5}.

(3) We chose the initial values of $\mathbf{p}_0$ and $M_{1,1}$ to be zero. The moments were chosen to be in the range $M_{2,0}\in[0.75, 3.0]^{[18]}$ and $M_{0,2} = dMk_BT$ where T is the desired temperature.

(4) At the beginning of the simulation, the system was integrated at constant temperature until it was judged to be at equilibrium. All measurements were taken after the equilibration period. The results for the static and dynamical properties for both linear and ring polymers using the MD and GPP methods are presented below.

B. Thermodynamic properties of the LJ homopolymers

An important measure of the polymer system is the gyration radius R_g defined as the root-mean-square distance between each monomer and the center of mass, e.g., for the GPP method

$$R_g^2 = \frac{1}{N}\sum_{k=1}^{N}[(\mathbf{r}_0^{(k)} - \mathbf{r}_c)^2 + M_{2,0}^{(k)}], \tag{13}$$

where N is the number of monomers in the chain, $\mathbf{r}_c$ is the center of mass of the polymer, and $\mathbf{r}_0^{(k)}$ and $M_{2,0}^{(k)}$ are the center and squared width of the packet representing the density of the kth site. For the MD method, the width term $M_{2,0}^{(k)}$ equals zero. Figure 1 shows the gyration radii as functions of temperature for both linear and ring 60mers using the MD and GPP methods. As expected, at high temperatures the R_g of the linear polymer exceeds that of the ring polymer. However, the value of R_g for both the linear and ring polymers converged to the same value in the low temperature limit. In the case of the GPP dynamics the high temperature limit of R_g is reached at a lower temperature than in the MD
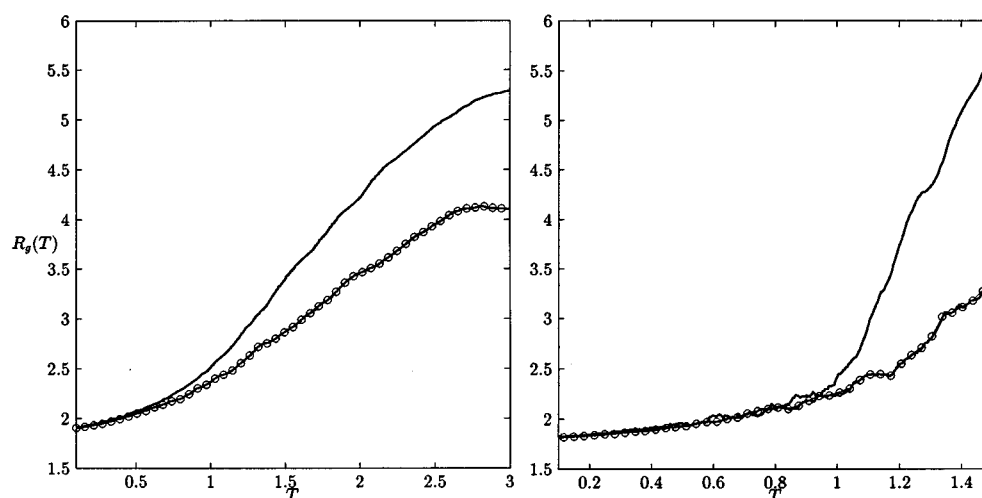
FIG. 1. The temperature dependence of gyration radii for 60mer linear (lines) and ring (circles) polymers. Results for the MD (left panel) and GPP (right panel) methods are shown. At low temperatures, the two kinds of polymer systems presented here have the same R_g values. The comparison suggests that the two methods give qualitatively similar results but differ by some temperature scale.

simulation. The comparison shown in Fig. 1 suggests that the MD and GPP results differ just by a scaling of T.

It is helpful to investigate the extent of the density fluctuations of the polymer chain in the course of transition. As a measure of these fluctuations, we calculated the distribution of gyration radii at different temperatures. Figure 2 shows typical distributions of gyration radii $P(R_g)$ at high, intermediate, and low temperature from the GPP simulation. At very high temperature, the polymer is in the random coil state. Not surprisingly, the R_g distribution is wide indicating that in the random coil state the density fluctuations are of the order of the density itself.[14] At intermediate temperature (close to transition) fluctuations are still pronounced. At lower temperature the distribution of R_g is a narrow peak indicating that the polymer molecule has a stable globular state at lower T. These distributions clearly demonstrate the existence of a coil-globule transition for the LJ homopolymer.

C. Dynamic properties of the LJ homopolymers

The kinetics of collapse of a polymer coil after an abrupt decrease in temperature by ΔT was investigated. A mechanism for this kinetic process was suggested by de Gennes[15] and then developed, to take into account topological constraints, by Grosberg, Nechaev, and Shakhnovich.[16] It was reported that one must decrease the temperature significantly by $\Delta T > \theta / N^{1/2}$ in order to have a significant collapse. Here θ is the temperature at which the second virial coefficient vanishes,[26] analogous to the Boyle temperature of gases.

Both MD and GPP simulations were run for 20–30 time units at higher temperature to calculate the average quantities in the coil state. This was followed by a one step sudden decrease in temperature, ΔT. In our simulation we chose the initial temperature of 1.5 and the final temperature of 0.6. We followed the system for a sufficiently long time after the quench to study the kinetics of the collapse of the polymer

chain. Several quantities, which are important for understanding the dynamics of the collapse of a LJ homopolymer, were measured.

We define a ''contact loop'' of the polymer chain as the loop whose head monomer i and end monomer j are within the defined contact distance δ. The distribution of the lengths of the contact loops (DCL) formed by the polymer chain is an important quantity which can be used to statistically describe the geometrical structure of a polymer. In our simulation, we chose the head-end contact distance $\delta = 1.3$ which is slightly larger than the equilibrium bond length. We traced the DCL as a function of time during the collapse of the chain. Figure 3 shows the representative pictures of the DCL for both linear and ring 60mers, calculated from GPP dynamics, after the temperature was suddenly decreased. The time evolution of the DCLs demonstrates that longer loops formed when the polymer chains relaxed following the sudden temperature quench. One can see from Fig. 3 that the number of contact loops shows a rapid initial increase and then becomes relatively stable in the long time limit. This reflects the fast geometrical structural collapse subsequent to the quench and the relatively stable globular state formed once the chain reaches equilibrium. The similarities in the DCLs of the linear and ring polymers in the long time limit means that the structures of the globular chains remain in a stable phase although they undergo slow internal structural changes. We observed similar dynamical features of DCLs for the conventional MD simulations.

In Fig. 4 we show the average length of the contact loop $\bar{L}$, calculated from DCLs, as a function of time for both linear and ring polymers. The results are presented for both MD and GPP dynamics. The $\bar{L}$ values have not reached a plateau even after the polymer molecules reach a globular state where the length of the contact loop should be rela-
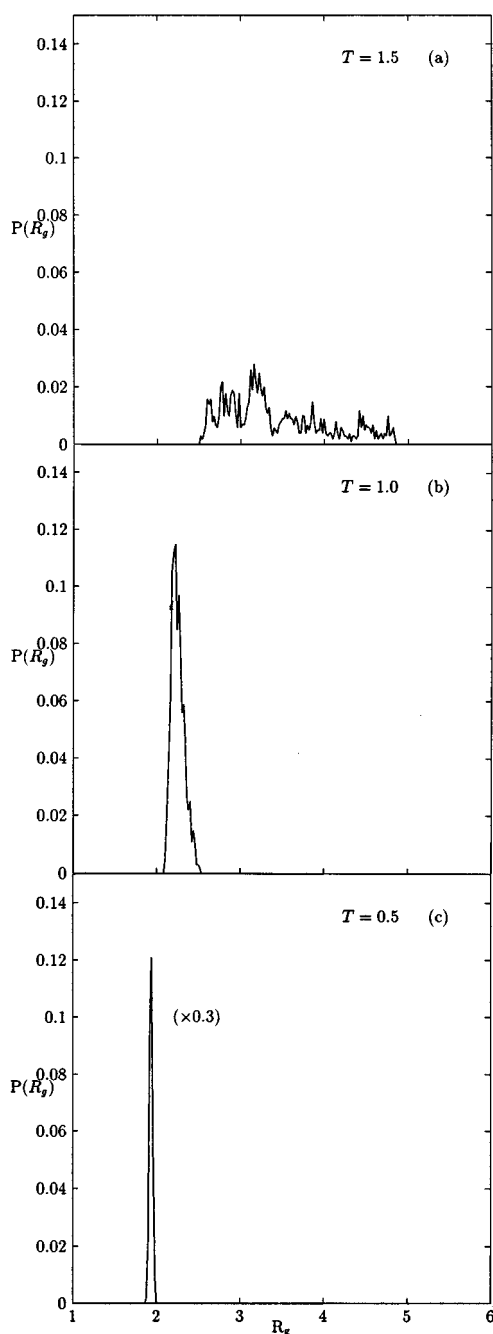
FIG. 2. The gyration radii distributions for the linear 60mer are presented at (a) high, (b) intermediate, and (c) low temperature, respectively. The results presented are from GPP simulations. At high T, the fluctuation in the polymer size is large while there is a well defined R_g at low temperature.

tively stable. Note that the average length of the contact loop for the ring polymer increased significantly less than that of the linear polymer. We believe this is due to the topological constraint of the ring polymer (where the head and end are joined). The slight positive slope in the curves in the long time limit suggests that the polymer chains are undergoing internal conformational rearrangement in the globular state. In the GPP simulation, the initial value of $\bar{L}$ is smaller than

that in the MD for both linear and ring polymers. However, the rate of increase of $\bar{L}$ in time is significantly larger in the GPP simulation indicating a faster reorganization.

We also present representative coordinate snapshots of the same relaxation process in Fig. 5 taken from GPP simulations. These snapshots support the conclusions drawn from other dynamical measures of the collapse process for the LJ homopolymer. In Fig. 6 we show the dependences of the gyration radii on time during the process of collapse. As one can see, the gyration radius converges to a stable value after a steep initial decay which implies that the polymer chains remain in a state with a stable density in the long time limit. This indicates the presence of a transition from the coil to the globular state of the polymer chain after the abrupt temperature drop. The final collapsed linear and ring polymers have very similar gyration radii. Note that the GPP simulations collapse the polymer (ring and linear) into a more compact state than in the MD simulation. It is also interesting to compare the gyration radii and the average length of a contact loop. We saw a slight increase in the average loop length at longer time while we do not see any significant change in R_g. This behavior may correspond to the second stage of relaxation for linear polymer chains predicted in the GNS scenario.[16] This stage corresponds to reorganization (potentially chain end reptation) which leads to formation of an equilibrium globule with a predominance of long range (along the sequence) contacts.

Thus, the simulations appear to support the theoretical prediction of a two-stage collapse of a homopolymer.[16] In the first stage, the polymer chain presumably collapses to a "crumpled" state; in the second stage, the crumpled state reorganizes further to a more globular state. It is not easy to fully characterize the crumpled stage for our system; however, we can recognize the continuous internal movement of the polymer chain in the globular state.

In order to study the structural properties of our chains, we also define a quantity R_s as a real space distance between two monomers which are g monomers apart along the polymer chain. The average value $\bar{R}_s(g)$ is defined as

$$\bar{R}_s(g) = \frac{1}{N-g} \sum_{i=1}^{N-g} |\mathbf{r}_0^{(i)} - \mathbf{r}_0^{(i+g)}|, \tag{14}$$

where N is the number of monomers. We calculated the value of $\bar{R}_s$ during the collapse of the chain. The results for $\bar{R}_s$ at different time intervals are shown in Fig. 7. The downward shifting coincides with a decrease in the physical size of the polymer molecule during the collapse.

It is of particular interest to study the short length lobe of the $\bar{R}_s$ curve. We anticipate that the short length lobe takes the form g^α where the value of factor α can be obtained by fitting the logarithm of the $\bar{R}_s$ curves. Theoretically, for a Gaussian-like chain[16] $\bar{R}_s(g) \propto g^{1/2}$, while for the crumpled globular conformation $\bar{R}_s(g) \propto g^{1/3}$. For the equilibrium globule, $\bar{R}_s(g) \sim g^{1/2}$ for $g < g_c$ and $\bar{R}_s \propto g^0$ for $g > g_c$. Here the crossover length g_c is determined by the condition that the size of a Gaussian fragment of the chain is of the order of the globule size R_g, i.e., $g_c \sim (R_g/a)^2$ where a is the average bond length. The time dependence of the exponent α for
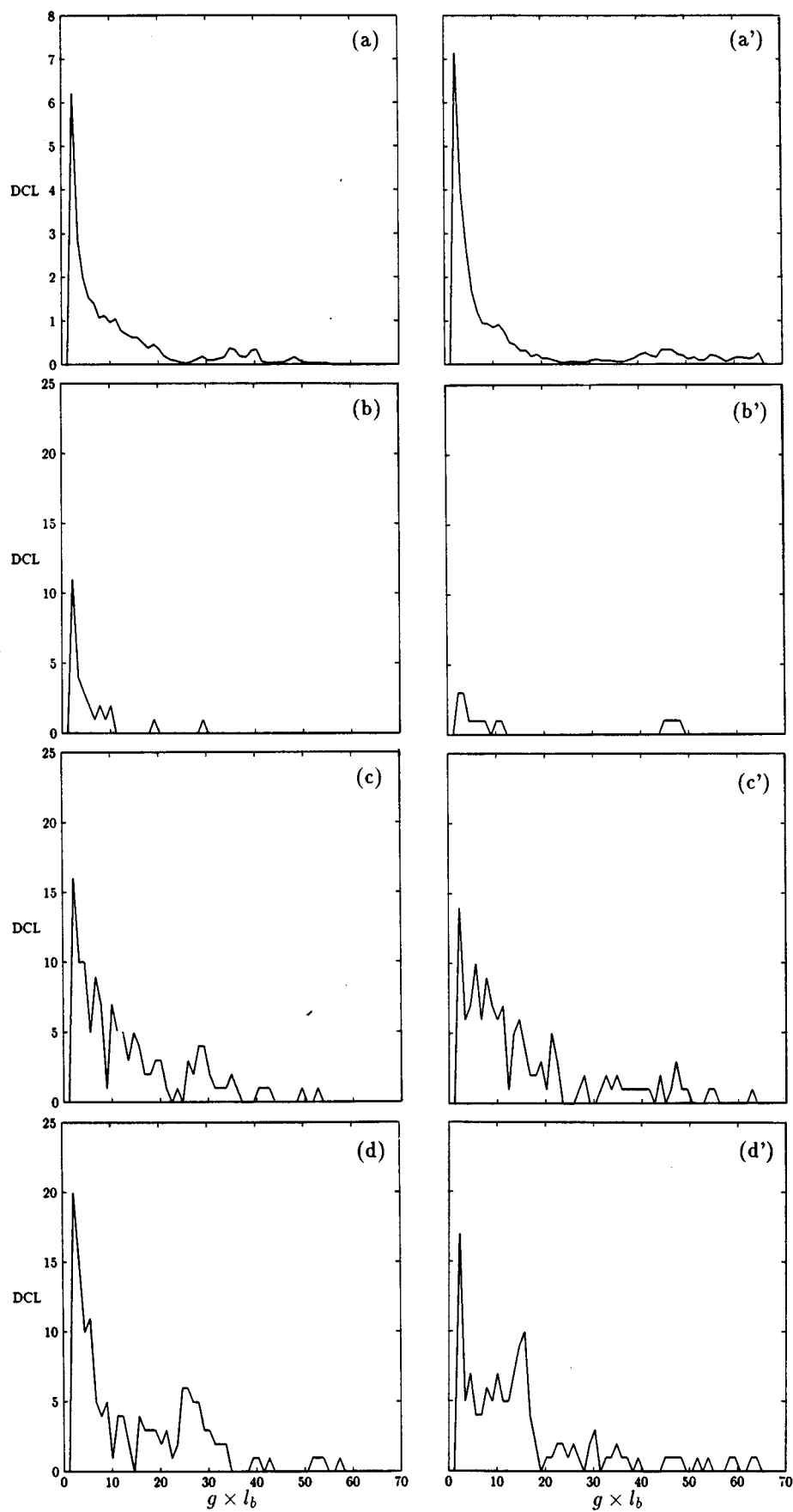
FIG. 3. Time dependence of the distribution of contact loops (DCL) for both linear cand ring polymers. GPP dynamics results are shown. The left panel column is for the linear 60mer: (a) High temperature average while (b), (c), (d) are typical instantaneous snapshots. The right panel column is for the ring 60mer: (a′) High temperature average while (b′), (c′), (d′) are typical snapshots.
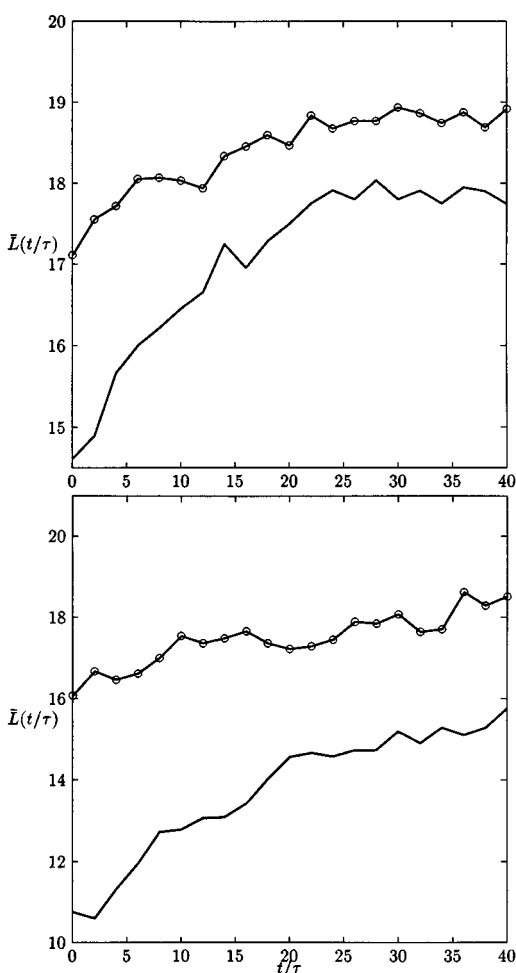
FIG. 4. Time dependence of the average length $\bar{L}$ of the contact loop for linear (lines) and ring (circles) molecules. Results are shown for both MD (upper panel) and GPP (lower panel) dynamics. The curves show that the length of the contact loop increases significantly during the collapse for the linear polymer while the ring polymer shows a relatively small change.

the collapse processes are shown in Fig. 8. The α values start around 1/2 in the high temperature limit and decay quickly following the temperature quench. This behavior is consistent with the results for the time dependence of the gyration radii. In the high temperature regime we can consider the polymer to be an ideal chain. The α values derived from simulation for the linear and ring molecules are in reasonable agreement with the theoretical predictions. At long times, α asymptotically reaches its equilibrium value. From the simulations, we found the α value at the final stage to be 1/3 for the ring polymer and a smaller value for the linear molecule. The simulation data give the final value of $\alpha = 1/3$ for the ring polymer which supports the assertion[16] that such polymers are in the crumpled globular state. For the linear polymer case, we see a smaller α value which again supports the theoretical prediction that the linear polymer molecule can reach an equilibrium globular state through a slow head-end rearrangement, in which case $\bar{R}_s$ is independent of α. It might be helpful to see at what length of the chain segment

the $\bar{R}_s(g) \propto g^\alpha$ relation breaks down. We found this relation to be reasonably accurate for chain segments of length less than 20 for the linear and ring polymers. Again, this is due to the fact that for large chain segments, the fragment size becomes comparable with the size of the globule as a whole.

Our results show that the α values obtained from GPP dynamics are always smaller than those of the MD results in the long time limit. We believe that this is due to the fact that, at finite temperature, the potential surface smoothing effect of the GPPs allows the polymer chains to relax more readily than those simulated by MD. This interpretation is consistent with the fact the equilibrium values of R_g derived from GPP dynamics are always slightly smaller than those of MD. These conclusions support the notion that GPP dynamics provides enhanced phase space sampling.[18,23]

IV. CONCLUSION AND DISCUSSION

This numerical simulation study of the thermodynamic and kinetic properties of Lennard-Jones homopolymers and the comparison with theoretical predictions can be concluded as follows:

(1) From the simulation study of the thermodynamic properties of a LJ homopolymer, we found that there does exist a clear coil-globule phase transition of the sort theoretically predicted by Lifshitz.[27] This phase transition appears to be second order.

(2) The process of collapse of the linear LJ homopolymer is a two-stage process, in accord with the theoretical prediction made by GNS.[16] In the first stage, the polymer becomes compact. In the second stage, the subsequent relaxation appears to involve the migration of chain ends and is correlated with changes in the topological structure of the chain.

(3) The structural properties of the linear and ring polymers are different; the ring polymer exhibits certain features of the crumpled globule.

(4) The dynamical mechanism of LJ homopolymer collapse does not appear to involve an activation barrier.

(5) Comparison of approximate GPP dynamics with exact MD simulations shows that GPP simulations provide a reasonably accurate representation of the system thermodynamics and dynamics. Differences arise in the rate of relaxation at low temperatures indicating the GPP dynamics are more effective in allowing reorganization in the compact state.

Simulation of the coil-globular transition was the subject of several previous studies.[11,28,29] Most of these studies employed MC simulation techniques in lattice models. The possible two-stage character of the collapse transition was noted by Kron et al. as early as in 1968. In recent work by Chan and Dill,[11] a two-stage collapse transition was observed at very low temperature. However, the origins of the stages observed by Chan and Dill and the one reported in this paper, as well as predicted by the GNS scenario, are quite different. The second stage which Chan and Dill observed only at very low temperature was found to be dependent on the move set used in the Monte Carlo simulation and is likely to be connected with final, fine arrangement of loops on the square
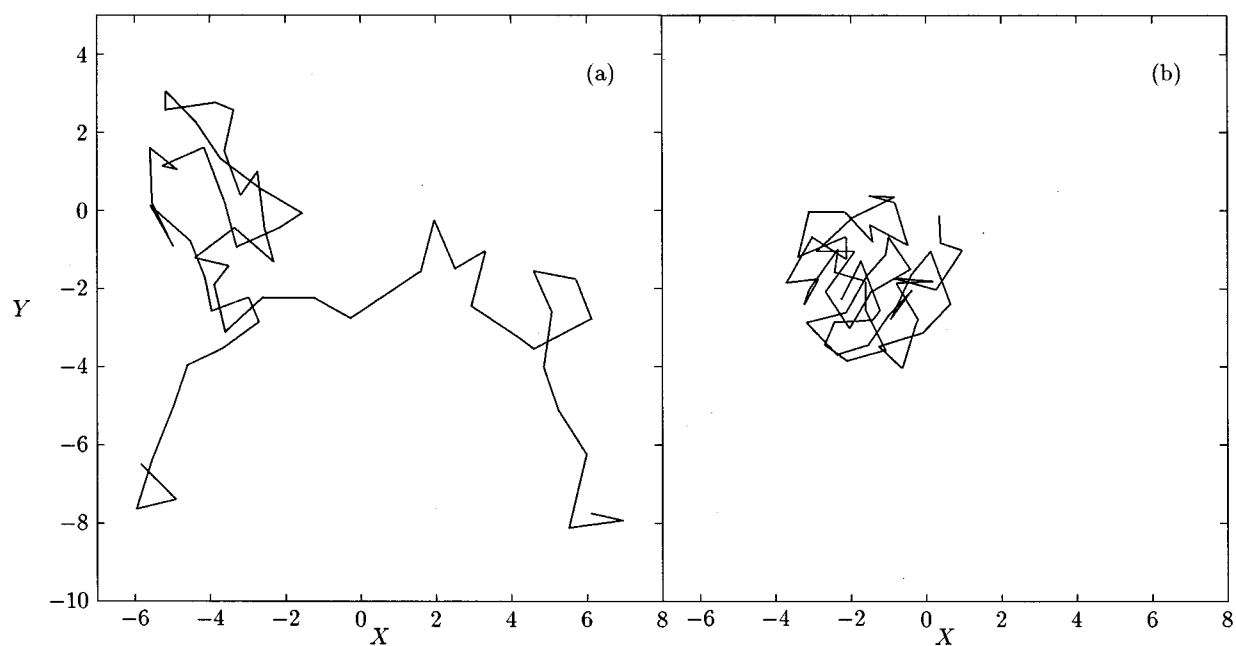
FIG. 5. Coordinate snapshots for the collapse of the linear 60mer. Results are from GPP dynamics. The graphs show the XY projection for (a) the initial snapshot at high temperature and (b) the final collapsed globular structure. The pictures show the clear collapse of the polymer molecule.

lattice. It should be contrasted with the second stage of the GNS scenario—a topological relaxation process which is not possible in the two-dimensional case studied by Chan and Dill.[11]

As we emphasized in the Introduction, many features of protein folding may be understood from polymer principles and statistical mechanics. The first stage of the folding process, which has been observed in many proteins,[2,4,30] is a fast collapse to a compact but nonspecific state. This bears resemblance to homopolymer collapse, and our aim in this
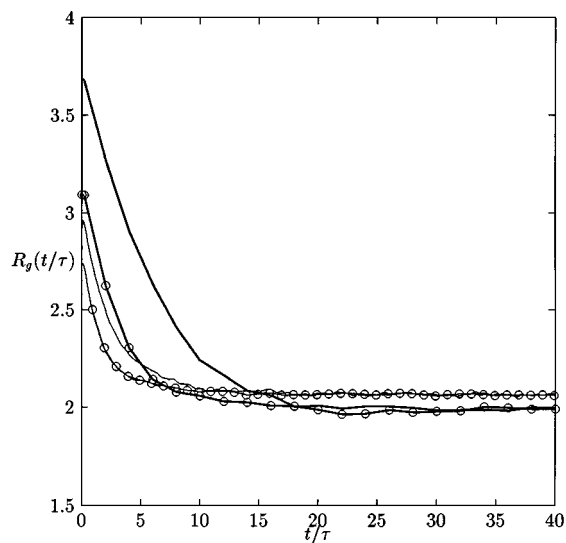

FIG. 6. The time dependence of the gyration radii during collapse. The curves record the rapid decrease in R_g in the early stage of collapse before relaxing to the same value of R_g for both linear (lines) and ring (circles) molecules at the end of the simulation. Results are shown for MD (thin lines) and GPP (thick lines) dynamics.
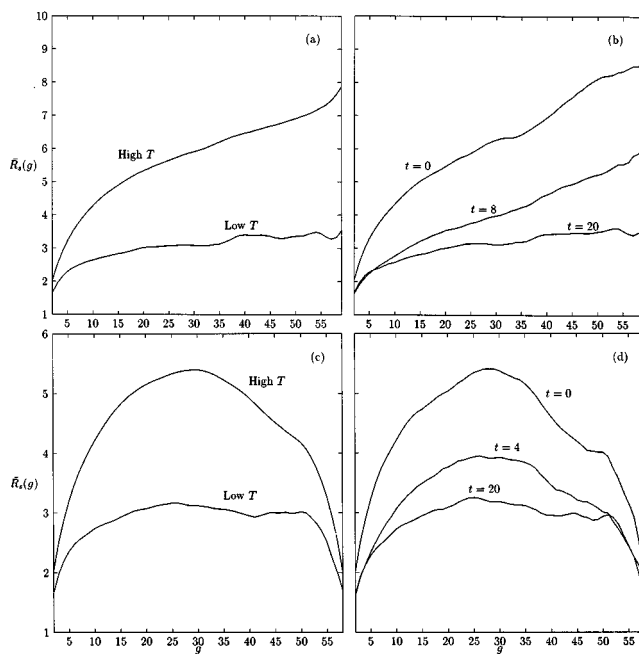

FIG. 7. Time dependence of $\bar{R}_s$ for the linear and ring molecules. Linear: (a) High and low temperature average and (b) intermediate time snapshots. Ring: (c) high and low temperature average and (d) intermediate time snapshots. GPP dynamics results are shown.
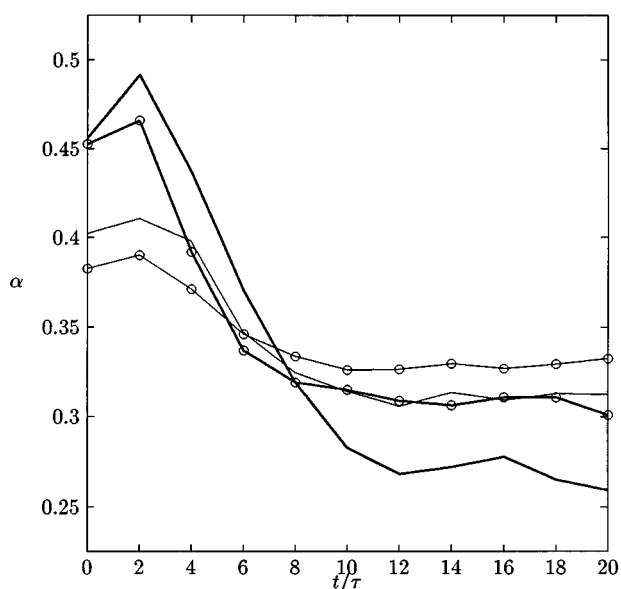
FIG. 8. The time dependence of α for the linear (lines) and ring (circles) polymers. The curves start from the high temperature average values of α, around 1/2 for both polymers, at $t=0$ and follow the collapse in time. At the end of the collapse (low temperature) α is equal to 1/3 for the ring polymer and a smaller value for the linear polymer. Results are shown for both MD (thin lines) and GPP (thick lines) dynamics.

FIG. 9. Bond length distributions are presented for three different temperatures. The equilibrium length of each bond is equal to $2^{1/6}$. The calculations are presented for a linear 60mer. Similar results can be expected for all other systems studied.

paper was to elucidate its possible kinetic mechanism. Subsequent stages of folding include the formation of specific unique structures. In these stages, heteropolymeric effects are of key importance. Therefore, a natural extension of the present study is to consider heteropolymers using the same simulation methodology. An important element of such studies would be sequence design, since random heteropolymers are unlikely to fold into the stable native state, for kinetic reasons.[5,31] This is a subject for future research.

ACKNOWLEDGMENTS

APPENDIX

The approximate density distribution used to describe the LJ homopolymers in our simulation is discussed in detail here. The bond potential is a part of the total potential surface, and when we calculate the average potential $\langle V \rangle$ it is most natural to average the harmonic bond potential over the density distribution. Doing that results in a term in $\nabla_{\mathbf{r}_0}^2 \langle V \rangle$ which is proportional to the bond force constant κ. This term dominates $\nabla_{\mathbf{r}_0}^2 \langle V \rangle$ making $\dot{M}_{2,0}$ large and negative. The result is that the widths of the packets shrink too fast and consequently the size of the polymer chain will be significantly smaller in the high temperature limit. Hence, we
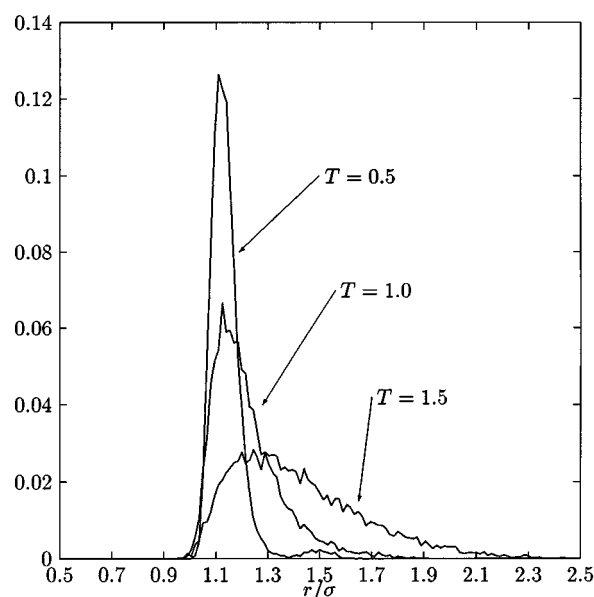
adopted an alternative approach. The bond potential averaged over the density distribution was approximated *ad hoc* as

$$\langle V \rangle_{\text{bond}} = \kappa \sum_{k=1}^{N-1} (|\mathbf{r}_k - \mathbf{r}_{k+1}| - l_b)^2$$

$$\cong \kappa \sum_{k=1}^{N-1} (|\mathbf{r}_{k,0} - \mathbf{r}_{k+1,0}| - l_b)^2. \quad \text{(A1)}$$

As such, there is no contribution from $\langle V \rangle_{\text{bond}}$ in $\nabla_{\mathbf{r}_0}^2 \langle V \rangle$. To test the properties of this representation, we investigated the bond fluctuation of the polymer chain. We calculated the bond length distribution at different temperatures. Figure 9 shows a typical bond length distribution at three temperatures. The fluctuations are relatively large at higher temperatures with smaller deviations at lower temperatures.

A standard result from polymer physics is that at high temperatures the gyration radius $R_g = \sqrt{l_b^2/6} N^{1/2}$ where l_b is the length of the bond. We calculated the gyration radius as a function of the chain length N. Figure 10 shows the size dependence of the gyration radius for a broad range of polymer sizes. In calculating the R_g the widths contribute significantly. Therefore, if Eq. (A1) is a poor approximation, we would expect to see an improper scaling of R_g. In Fig. 10, the linear fitting of the data shows that the slope is 0.48 which is very close to 1/2. We also calculated the value of l_b from the extrapolation of Fig. 10 to $N=1$ which provides a value of the bond length equal to 1.48 in the high temperature limit at $T=1.5$. Meanwhile, we compared this bond length with the average bond length calculated from the distribution shown in Fig. 9 which is equal to 1.41 at $T=1.5$.
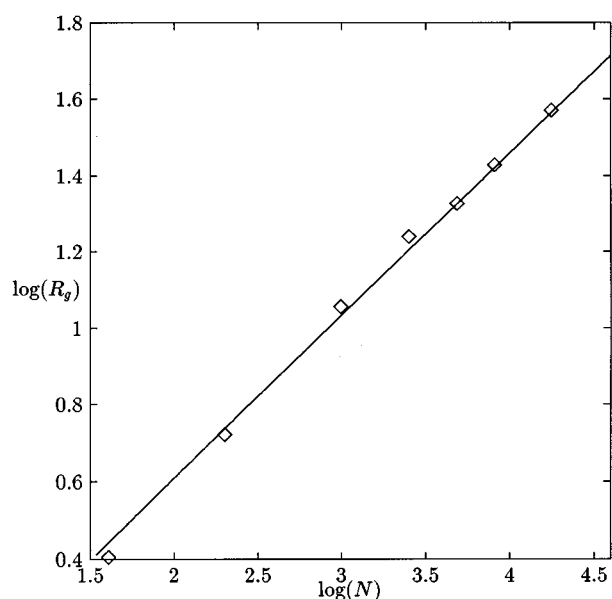
FIG. 10. Polymer size dependence of the gyration radius. The results are shown in the logarithmic scale for a linear polymer in the high temperature limit $T = 1.5$. The diamonds are for the simulation data and the solid line is for the linear fit.

So the expected bond length from theory agrees with the average bond length from simulation. Therefore, we argue that, in our simulation, the Gaussian ideal chain is well represented by the approximate phase space density distribution we used.

Also, for comparison with the well depth of the non-bonded Lennard-Jones potential equal to one, the curvature of the LJ potential at its minimum is 36 in reduced units. However, the bond potential between sites on the homopolymer is not intended to mimic a chemical bond. It is taken as a reasonable estimate of the fluctuations between centers of polymeric subunits along the polymer chain. As a result, the bond force constant is weak compared with a typical covalent bond to allow for the larger fluctuations between the centers of the subunits, which has contributions from bond, angle, and torsional fluctuations. We also note that if each site is considered to be a subunit (such as an amino acid)

where the diameter is 10 Å, the mass 100 gm/mol, and the well-depth 3 kcal/mol, the time unit is approximately 150 ps.

[1] H. Frauenfelder and P. G. Wolynes, *Physics Today*, Feburary 1994.
[2] For a recent review of experimental studies of protein folding, see R. Baldwin, Curr. Opinion Struct. Biol. **3**, 84 (1993).
[3] A recent review of theoretical approaches to the protein folding problem is given by M. Karplus and E. Shakhnovich, in *Protein Folding*, edited by T. Creighton (Freeman, San Francisco, 1992).
[4] S. Radford, C. Dobson, and P. Evans, Nature **358**, 302 (1992).
[5] E. I. Shakhnovich, G. M. Farztdinov, A. M. Gutin, and M. Karplus, Phys. Rev. Lett. **67**, 1665 (1991).
[6] P. E. Leopold, M. Montal, and J. N. Onuchic, Proc. Natl. Acad. Sci. U.S.A. **89**, 8721 (1992).
[7] C. Camacho and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **90**, 6369 (1993); D. Thirumalai and Z. Guo, Biopolymers **35**, 137 (1995).
[8] A. Sali, E. I. Shakhnovich, and M. Karplus, Nature **369**, 248 (1994).
[9] N. Socci and J. Onuchic, J. Chem. Phys. **101**, 1519 (1994).
[10] H. S. Chan and K. A. Dill, J. Chem. Phys. **100**, 9238 (1994).
[11] H. S. Chan and K. A. Dill, J. Chem. Phys. **99**, 2116 (1993).
[12] O. B. Ptitsyn and Y. E. Eisner, Biofizika **10**, 3 (1965).
[13] P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University, Ithaca, NY, 1979).
[14] I. M. Lifshitz, Sov. Phys. JETP **28**, 1280 (1969); Zh. Eksp. Teor. Fiz. **68**, 2408 (1968).
[15] P. G. de Gennes, J. Phys. Lett. **46**, L-639 (1985).
[16] A. Y. Grosberg, S. K. Nechaev, and E. I. Shakhnovich, J. Phys. (France) **49**, 2095 (1988).
[17] M. Doi and S. F. Edwards, J. Chem. Soc. Faraday Trans. 2 **74**, 1789, 1802, 1818 (1978).
[18] J. Ma, D. Hsu, and J. E. Straub, J. Chem. Phys. **99**, 4024 (1993).
[19] R. C. Tolman, *The Principles of Statistical Mechanics* (Dover, New York, 1979).
[20] B. J. Berne and R. Pecora, *Dynamic Light Scattering* (Wiley-Interscience, New York, 1976).
[21] R. W. Zwanzig, in *Lectures in Theoretical Physics*, edited by W. E. Britton, B. W. Downs, and J. Downs (Wiley-Interscience, New York, 1961), Vol. III, p. 106; see also J. Chem. Phys. **33**, 1338 (1960).
[22] W. G. Hoover, A. J. C. Ladd, and B. Moran, Phys. Rev. Lett. **48**, 1818 (1982); D. J. Evans, J. Chem. Phys. **78**, 3297 (1983).
[23] J. Ma and J. E. Straub, J. Chem. Phys. **101**, 533 (1994).
[24] P. Amara, D. Hsu, and J. E. Straub, J. Phys. Chem. **97**, 6715 (1993).
[25] S. A. Teukolsky, W. H. Press, B. P. Flannery, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1986).
[26] P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, NY, 1953).
[27] A. Yu. Grosberg and A. R. Khokhlov, Sov. Sci. Rev. A. Phys. **8**, 147 (1987).
[28] A. Kron, O. B. Ptitsyn, and Y. E. Eisner, J. Polym. Sci. **16**, 3509 (1968).
[29] B. Ostrovsky and Y. Bar-Yam, Europhys. Lett. **25**, 409 (1994).
[30] S. Khorasanizadeh, I. D. Peters, T. R. Butt, and H. Roder, Biochemistry **32**, 7054 (1993).
[31] E. I. Shakhnovich and A. M. Gutin, Proc. Natl. Acad. Sci. U.S.A. **90**, 7195 (1993).